

Paradigmenwechsel bei Wissensvermittlung wegen dynamischer Veränderungsprozesse¹

Andreas Degkwitz – UB der Humboldt-Universität zu Berlin

Wissenskulturen, wie sie auch von Bibliotheken repräsentiert werden, stehen einerseits mit Kulturtechniken und andererseits mit dem darauf aufsetzenden Wissensverständnis in Zusammenhang – vor folgendem Hintergrund: Schon immer haben Menschen sich darum bemüht, große Herausforderungen an Technik auszulagern. Ein sehr gutes Beispiel dafür ist das Automobil, an das vor allem körperliche Bewegung von A nach B ausgelagert wird. Seit einigen Jahrtausenden lagern wir unser Gedächtnis zunächst an die Technik der Schrift, dann an beschriftete Trägermedien (Papyrus, Pergament, Papier) und nun an Computer, Netze und Software aus. Indem wir Verarbeitung (Processing), Verbreitung (Distribution) und Speicher (Memory) auf Server und Netze auslagern, sind wir in der Situation, in einem Umfang auszulagern, der uns bisher nicht zur Verfügung stand: Wir konnten noch nie so viel an Gedächtnis, Verarbeitung und Verbreitung auslagern wie es uns heute möglich ist. Aber wir laufen auch Gefahr, durch Auslagerung zu verlieren, so dass sich die Frage stellt: Was machen wir mit diesen neuen Maschinen und was machen diese Maschinen mit uns?

Aus einer rein technischen Sicht verstehen sich Texte als Inhalte in geschriebener Sprache, die durch eine lineare Abfolge alphanumerischer Zeichen konstituiert sind. Ein solches Textverständnis ist wesentlich für darauf aufsetzende, technische Arbeitsschritte zur IT-gestützten Verarbeitung von Texten. Der diesen Verfahren zugrunde liegende, binäre Code, der die alphanumerischen Zeichen konstituiert, legt die Annahme nahe, dass es dabei weniger um die Verarbeitung von Texten als viel-

¹ Im Rahmen der Ringvorlesung des Departments „Wissen – Kultur – Transformation“ der Interdisziplinären Fakultät der Universität Rostock „Wissen: Ideal – Kultur – Institution. Ringvorlesung zur Wissens-, Wissenschafts- und Universitätsgeschichte“ wurde am 8. Mai 2012 eine erweiterte, nicht veröffentlichte Fassung des vorliegenden Beitrags vorgetragen. s. <http://www.inf.uni-rostock.de/departments/wissen-kultur-transformation/veranstaltungen/>. Eine Kurzversion dazu ist in Bibliotheksdienst 46 (2012) unter dem Titel „Linked Knowledge – Open Libraries“ (S. 570 – 578) erschienen. Eine wissenschaftlich aufgearbeitete Fassung des Beitrages ist mit dem Titel „Texte, Daten, Bilder – Wissen!“ in Bibliothek, Forschung und Praxis 36 (2012) veröffentlicht (S. 215 – 219).

mehr um die Verarbeitung von Daten geht. Denn dem Code des Alphabets wird ein Code zugrunde gelegt, der die Schrift als Abfolge binär codierter Buchstaben abbildet oder simuliert. Die Umsetzung des Zeichencodes in Schrift und in den sich daraus ergebenden Text ist nichts anderes als die Konfiguration oder ein „Ausgabeformat“ von Daten, die nur deshalb als alphanumerische Zeichenfolge umgesetzt sind, weil wir diese bis dato besser und schneller aufnehmen und rezipieren können als alle anderen grundsätzlich denkbaren Ausgabeformate, die der Binär-Code editieren kann.

Der maschinell erstellte Text ist von daher eine Abfolge binär codierter Daten, wie es der am Computer lesbare Text ebenfalls ist. Dabei werden Daten als Zeichen oder Zeichenfolgen mit – im Regelfall sehr unterschiedlichen und sehr unterschiedlich ausgeprägten – Strukturinformationen (Metadaten) verstanden. Deutlich wird das gerade auch bei Texten, die maschinell verarbeitet werden können und deshalb z.B. als HTML- oder XML-File strukturiert sind. Zwar greift die Strukturierung nach HTML oder XML die aus der Printwelt bekannten Strukturierungsinformationen auf, da offenbar bisher keine besseren Strukturierungsmodelle bestehen. Doch im Ergebnis konvertiert dieser Prozess den Text zum Datenpool oder Datencontainer, den die weitere maschinelle Verarbeitung, Verbreitung und Speicherung erforderlich macht. Die Repräsentation von Wissen erfolgt damit in Daten und Datenbestände, die in unterschiedlichsten Formaten mit sehr heterogenen Provenienzen auf Computern und in Netzen verfügbar sind.

Ob der Wandel vom Pergament zum Papier, von der Handschrift zum Buchdruck, von der Schrift zum Text ein größerer, weitreichenderer war, als der von Druckwerken zu Maschinen und von Texten zu Daten? Gravierend ist der Medienwandel auf jeden Fall und ist ähnlich den Etappen, die ihm vorausgingen, mit einem neuen Wissensverständnis verbunden; das konfrontiert uns mit dem Thema der Informationsflut, die eigentlich eine Datenflut ist und uns als Informations- und Wissensgesellschaft charakterisiert. Dabei meinen wir in einer Wissens- und Informationsgesellschaft zu leben und müssen uns immer wieder aufs Neue unser Nicht-Wissen eingestehen. Erklären lässt sich dies vielleicht so, dass wir einerseits ei-

nen schier unbegrenzten Zugang zu „Wissen“ haben, dass wir uns aber andererseits kaum in der Lage sehen, dieses „Wissen“ zu rezipieren. Wahrscheinlich verstehen wir unser Unvermögen besser, wenn wir nicht von Wissen, sondern mehr von Informationen und noch besser von Daten sprechen, deren Fülle zu verarbeiten uns in jeglicher Hinsicht überfordert.

An dieser Stelle hilft uns das der neuen Kulturtechnik innewohnende Potenzial, Gedächtnis- und Verarbeitungsleistungen in einem bisher ungekannten Umfang auslagern zu können. So übertragen wir die Verarbeitung von Daten sowie deren Recherche und Speicherung an Server, Software und Netze. In der weiteren Folge stellt sich die Frage, zu welchem Verständnis von Wissen diese Kulturtechnik führt. Denn bei dieser Repräsentation des Wissens geht es offenbar weniger darum, theoriebasiertes oder hermeneutisches Wissen hervorzubringen. Vielmehr führen die datengestützten Verfahren zu empirischer Forschung, für die Analyse und Auswertung ermittelter Daten kennzeichnend ist. Mit Hilfe von Maschinen (Endgeräte, Netze, Software und Server) werden Daten recherchiert, aggregiert, prozessiert und evaluiert, um auf diese Weise Aufschlüsse und Antworten zu empirischen Fragestellungen zu gewinnen.

Dabei geht es über die Fragestellungen der klassischen Erfahrungswissenschaften (Medizin, Natur- und Ingenieurwissenschaften) hinaus verstärkt auch um geistes-, kultur- und sozialwissenschaftliche Forschungen, die auf diesem Wege historische Zusammenhänge rekonstruieren, aktuelle Entwicklungen analysieren und künftige Szenarien prognostizieren. Deutlich wird, dass sich so neue Forschungsfelder eröffnen, die mit den vorausgegangenen Kulturtechniken nicht hätten bearbeitet werden können: Linguistische Forschungen können durch Ton- und Textdokumente signifikant ergänzt werden. In den Kunstwissenschaften werden epochen- und standortübergreifende Bildvergleiche möglich. Die Simulation von Artefakten kann Gemälde, Gebäude, Plastiken oder Skulpturen virtuell rekonstruieren. In den Geschichtswissenschaften lassen sich mit digitalisierten Urkunden, Dokumenten, Bildern, Büchern, Gegenständen und Filmen historische Entwicklungen in allen Lebensbereichen vielfältig belegen und veranschaulichen. Damit hat sich das Wissensverständnis zu einer datengestützten Empirie gewandelt, die auch in Fachdisziplinen, die

bisher eher theorieorientiert oder hermeneutisch agierten, den textbezogenen, spekulativen Ansatz von Theorie in den Hintergrund drängt. Das in Servern und Netzen repräsentierte, empirische Wissen ist ein Ergebnis maschinengestützter Filter- und Selektionsverfahren, deren inhaltliche Parameter am Frontend vom Nutzer (nicht Leser) gesetzt werden. Die neuen Medien begründen damit die aus Daten gewonnene Empirie als neues, aktuelles Wissensparadigma.

Mit Blick auf die skizzierten Entwicklungen stellt sich die weitere Frage, welches Format die Daten der Empirie in einen Zusammenhang bringt, der uns die Daten verstehen lässt. Wie lassen sich die Wissensergebnisse und die sie begründenden Datenbestände in ihrer Komplexität rezipieren? Diese Funktion hat für die schrift- und buchgebundene Aggregation der Text. Daten hingegen können weder aggregiert noch analysiert oder ausgewertet werden, wenn sie nicht „ins Bild“ gesetzt werden - genauer gesagt, in einen Kontext gesetzt werden, der, wie es Bilder vermögen, Komplexität reduziert, indem Komplexität veranschaulicht wird. Die Informationstechnik hat vielfältige Möglichkeiten, die Aggregation von Datenbeständen in dieser Weise zu präsentieren - mit anderen Worten: zu visualisieren oder zu simulieren. Der Kontext, der Schrift und Text auf dem Gebiet maschinenlesbarer Daten ersetzt, ist in einem sehr umfassenden Sinne das **‚Bild‘**, das ein technisches Bild ist und uns in Baumstrukturen, Buchstabenfolgen, Diagrammen, Grafiken, Präsentationen und in welcher Gestaltung auch immer – durchaus auch bewegt und dreidimensional – begegnet: Das Spektrum reicht von Tortendiagrammen für statistische Erhebungen über Ablauf- und Organisationsmodelle bis hin zu komplexen Simulationen, die datenbasierte Kontexte visualisieren, die ohne diese Form der Veranschaulichung weder erkennbar noch nachzuvollziehen wären.

Ähnlich den ‚Images‘ für Computerinstallationen, die sowohl Informationen als auch Strukturen des File-Systems umfassen, werden künftige Publikationen - als digitale Objekte - nicht nur Daten, sondern auch Kontexte und Strukturen von Datenbeständen enthalten. Mit anderen Worten: Wenn wir Ergebnisse unserer empirischen Forschungen präsentieren wollen, die auf aggregierten, evaluierten oder prozessierten Daten beru-

hen, werden wir unsere Vorstellung von der Realität oder das, was wir für Realität halten, veranschaulichen oder simulieren. Insofern müssen Daten ‚ins Bild‘ gesetzt werden, um wahrgenommen und verstanden werden zu können. Der Kontext für die Aggregation von Daten ist das Bild – allerdings ein Bild, das auf der Basis der ihm zugrunde liegenden Daten interoperabel ist und deshalb verarbeitet und verändert werden kann.

Der prozessuale Charakter computergestützten Arbeitens hat zur Folge, dass über die globalen Möglichkeiten der Verbreitung von Inhalten hinaus Wissen und Wissensaneignung verstärkt als dynamische Prozesse aufgefasst und gelebt werden. Statische Wissensformen, wie sie vor allem durch gedruckte Veröffentlichungsformen repräsentiert werden, verlieren hingegen ihren bisherigen Stellenwert. Ausschließlich textbasiertes Wissen und seine gedruckten Präsentationen verlieren damit an Bedeutung. Stattdessen entstehen verstärkt datengestützte Publikationsformen, die multimedial angereichert und in beliebige Kollaborationskontexte eingebracht werden können oder auf kontextualisierten, verlinkten Datenbeständen unterschiedlicher Provenienz gründen.

In diesen Zusammenhang gehören beispielsweise auch Druckbestände, die nach OCR als Volltexte digitalisiert sind und für ihre Weiterverarbeitung nach RDF strukturiert werden; mit Hilfe von Suchinstrumenten lassen sich diese – auch als Ontologien - aggregieren und bieten damit die Grundlage für die Bildung von Kontexten zu den jeweils verfolgten Fragestellungen. Darüber hinaus bietet die Integration von Softwarekomponenten wie Apps die Möglichkeit, datenbasiertes Wissen unter vielen Aspekten und Fragestellungen unmittelbar zu verarbeiten. Unterschiedliche Informationsmaterialien können so in **einem** digitalen Dokument konvergieren und damit Bestandteile **einer** Publikation bilden oder sie werden als differenziert verlinkte Datenbestände veröffentlicht. Die neuen, digitalen Medienformate übersteigen damit das Potential gedruckter Formate signifikant. Dies hat Auswirkungen auf die Vermittlung und Verbreitung künftigen Wissens.

Insgesamt zeichnen sich als Grundanforderungen an ‚future publications‘ eine Aufbereitung der Publikationen nach folgenden Kriterien ab:

- Ergonomische Navigation für Präsentation und Rezeption von Text- und Bilddaten sowie weiterer Datenbestände,
- Strukturierung nach RDF, um Publikationen bis auf die Ebene einzelner Elemente (data sets) einerseits recherchieren und andererseits kontextualisieren zu können,
- Ausgabeformate (z. B. XML) für Interoperabilität und Integration in fachliche Netze und Plattformen,
- Verarbeitbarkeit (processing) der Veröffentlichung insgesamt und ihrer einzelnen Komponenten (data sets) - z. B. mit Hilfe von Apps - bis hin zu Data-Mining,
- Vergabe von Persistent Identifiern (PID) und Metadaten, um die Zugänglichkeit der Datenbestände langfristig sicherzustellen,
- Optionen für ein Zugangsmanagement (Authentifizierungsverfahren) für nachträgliche Anreicherungen, Anmerkungen und Kommentare sowie Nachnutzungen mit Social-Media-Tools,
- Voraussetzungen für ein Mobile-Device-Management, um auf unterschiedlichen, mobilen Endgeräten gelesen werden zu können.

Mit den Anforderungen an digitale Veröffentlichungen sind zugleich neue Service- und Unterstützungsprozesse der Informationsinfrastruktureinrichtungen (Bibliotheken, Medien- und Rechenzentren sowie Verlage, IT-Provider) erforderlich. Zu den neuen Aufgaben gehören Interface- und WEB-Design, Maßnahmen zur Digitalisierung, Forschungs- und Primärdaten-Management, Betrieb und Vernetzung von Repositorien, Hosting und Speicherung von Applikationen und Daten sowie Geschäfts- und Verwertungsmodelle. Im Unterschied zu den bestehenden Services von Bibliotheken, die von ihrer Struktur her auf gedruckte Inhalte und Texte ausgerichtet sind, steht bei ‚future publications‘ der Umgang mit Daten im Mittelpunkt: Daten und Datenbestände werden erhoben, prozessiert,

kontextualisiert, strukturiert, verlinkt, gehostet, archiviert und nicht zuletzt recherchiert, um wieder aufgefunden und nachgenutzt werden zu können. Für Bibliotheken ist diese Entwicklung eine große Chance und eine große Herausforderung, die unbedingt aufzugreifen ist. Ihre Serviceportfolios müssen auf Forschungsprozesse und Forschungsergebnisse ausgerichtet werden, für die das Textparadigma nicht mehr im Mittelpunkt steht, sondern die vorrangig auf Daten mit entsprechenden Strukturinformationen beruhen. Diese Transformation konfrontiert uns mit vielen Veränderungen, die mit einem tief greifenden Wandel des Wissensverständnisses und der Dienste und Services der Informationsinfrastrukturen einhergehen.

Literaturhinweise:

Degkwitz, Andreas: Texte, Daten, Bilder – Wissen! – in: Bibliothek, Forschung und Praxis 36 (2012), S. 215 – 219

Fournier, Johannes: Zugang, Nachnutzung und Reproduzierbarkeit. Anmerkungen zur künftigen Ausrichtung einer wissenschaftsadäquaten Informationsinfrastruktur. – in: Bibliothek Forschung und Praxis. Band 36, Seiten 180 –188

Dudek, Sarah: Die Zukunft der Buchstaben in der alphanumerischen Gesellschaft. Text und Dokument unter digitalen Bedingungen. – in: Bibliothek, Forschung und Praxis 36 (2012), S. 189 – 199

Füssel, Stephan (hrsg.): Medienkonvergenz – Transdisziplinär. - Berlin, de Gruyter (2012), S. 1 – 6 (Einleitung)

Mittler, Elmar: Wissenschaftliche Forschung und Publikation im Netz. - in: Medienkonvergenz - Transdisziplinär. - hrsg. von Stephan Füssel. - Berlin, de Gruyter, (2012), S. 32 – 80

Neuroth, Heike; Strathmann, Stefan; Oßwald, Achim; Scheffel, Regine; Klump, Jens; Ludwig, Jens (Hg.): Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme. Boizenburg, Göttingen: Hülsbusch; Univ.-Verl. Göttingen (2012). Online verfügbar unter http://www.nestor.sub.uni-goettingen.de/bestandsaufnahme/nestor_lza_forschungsdaten_bestandsaufnahme.pdf